Geophys. J. Int. (2006) **167,** 238–252 doi: 10.1111/j.1365-246X.2006.03084.x

Global multiresolution models of surface wave propagation: comparing equivalently regularized Born and ray theoretical solutions

Lapo Boschi
E.T.H. Zürich, Switzerland. E-mail: lapo@erdw.ethz.ch

Accepted 2006 May 19. Received 2006 May 18; in original form 2004 October 18

SUMMARY

I invert a large set of teleseismic phase-anomaly observations, to derive tomographic maps of fundamental-mode surface wave phase velocity, first via ray theory, then accounting for finite-frequency effects through scattering theory, in the far-field approximation and neglecting mode coupling. I make use of a multiple-resolution pixel parametrization which, in the assumption of sufficient data coverage, should be adequate to represent strongly oscillatory Fréchet kernels. The parametrization is finer over North America, a region particularly well covered by the data. For each surface-wave mode where phase-anomaly observations are available, I derive a wide spectrum of plausible, differently damped solutions; I then conduct a trade-off analysis, and select as optimal solution model the one associated with the point of maximum curvature on the trade-off curve. I repeat this exercise in both theoretical frameworks, to find that selected scattering and ray theoretical phase-velocity maps are coincident in pattern, and differ only slightly in amplitude.

Key words: scattering theory, seismic resolution, tomography, upper mantle.

1 INTRODUCTION

Tomographic images of the Earth seen in the literature are generally based on the ray-theory, or JWKB approximation. In practice, the momentum equation is solved in the limit of infinite frequency, with an application of the WKB approximation (e.g. Bender & Orszag 1978); the problem is thus separated into transport and eikonal equations, and seismic rays can be introduced as the characteristics of the eikonal equation (Červeny 1985). It follows that a seismic measurement approximately depends only on the velocity structure encountered along the ray path; for example, a body wave traveltime anomaly can be written as a line integral of slowness perturbations along the ray path, with respect to the reference model, and a linear tomographic inverse problem can thus be set up (e.g. Boschi & Dziewonski 1999).

Away from this limit, that is, as frequency decreases, it gradually becomes meaningless to speak of seismic rays, and the dependence of seismic measurements on the Earth's structure away from the JWKB ray path becomes more important; as long as the JWKB approximation is still applied, it is then believed that resolution (the real velocity anomaly of smallest volume that can be properly mapped through tomography) is limited by the wavelength of the inverted seismic measurements.

Many authors have attempted to go beyond ray theory, and incorporate finite-frequency effects in their *forward* models of seismic wave propagation: proposed alternative approaches involve the summation of normal modes (Woodhouse & Girnius 1982; Clévédé. & Lognonné 1996; Capdeville 2005), often simplified via the far-field (asymptotic) approximation (Snieder 1986a, 1987; Snieder & Nolet 1987; Li & Tanimoto 1993; Li & Romanowicz 1995; Dahlen *et al.* 2000; Hung *et al.* 2000; Tanimoto 2003; Zhou *et al.* 2004; Yoshizawa & Kennett 2005), or the expensive numerical integration of the equations of motion (e.g. Komatitsch *et al.* 2002). On the other hand, few accounted for finite-frequency effects in the formulation of *inverse* problems: the only published attempts in global tomography, all based on various formulations of perturbation theory (scattering, or, as we will dub it hereafter, Born theory), are those of Snieder (1987, 1988), Ritzwoller *et al.* (2002) (where Fréchet kernels were replaced with boxcar or Gaussian functions), Spetzler *et al.* (2002) and Zhou *et al.* (2005) in the context of surface wave dispersion; Li & Romanowicz (1996) (and, later, other authors from the Berkeley group), whose NACT approximation has the merit of accounting for mode coupling, providing 2-D waveform sensitivity kernels; Montelli *et al.* (2004a,b), who inverted body wave traveltime observations made at relatively long period.

Montelli *et al.* (2004a) claimed to have mapped anomalies of high spatial frequency (ascending plumes in the mantle, characterized by low seismic velocity), that would not have been visible through traditional, ray theoretical tomography; the general character of their images, however, does not seem to be much perturbed by the introduction of the more sophisticated theory. Finite-frequency effects should have an even stronger impact on surface wave tomography, which is based on the longer-period (typically 30–150 s) component of seismograms, further away from the high frequency limit;

nevertheless, Spetzler *et al.*'s (2002, fig. 3) comparison of Born versus ray theoretical tomographic maps does not evidence any coherent pattern, making it hard to prove that the application of scattering theory leads to an effective improvement of the images.

Spetzler *et al.* (2002) made use of a tomographic parametrization of relatively low resolution: spherical harmonics up to degree 40 (1681 free parameters). It is possible that phase-velocity anomalies of high spatial frequency, that would have been retrieved, thanks to the Born-theory formulation, could not be properly described by the degree-40 harmonic parametrization; in other words, the resolution gained by the application of scattering theory could have been lost owing to the long spatial wavelength of the parametrization (Appendix A).

I repeat Spetzler *et al.*'s (2002) exercise, describing surface wave phase velocity in terms of a grid of approximately equal-area pixels, with a lateral extent of $3^\circ \times 3^\circ$ at the equator. In North America (70°E to 130°E, and 15°N to 60°N), a region particularly well sampled by seismic sources and stations, I replace the $3^\circ \times 3^\circ$ grid with a finer one, subdividing each pixel in nine smaller ones (lateral extent of $1^\circ \times 1^\circ$ at the equator); this parametrization (6720 free parameters) should be globally at least as accurate as that of Spetzler *et al.* (2002), and much better within the high-resolution region. The database I invert, updated from Ekström *et al.* (1997), includes ∼30 000 summary dispersion observations, which guarantees that the inverse problem is not underdetermined. As recently suggested by Godey *et al.* (2003), comparison of Born versus ray theoretical maps in this region should tell us more about the importance of finite-frequency effects in global tomography.

2 THEORY

The earliest exhaustive work on the application of single-scattering theory to surface wave propagation is that of Snieder (1987); here I will give a brief, and deliberately, simplistic outline of his treatment.

Green's problem (e.g. Dahlen & Tromp 1998, section 4.1.7) associated with the Earth's elastic oscillations following an earthquake can be written as

$$L\boldsymbol{G} = \boldsymbol{F}, \tag{1}$$

where $\boldsymbol{G}$ is Green's tensor, $\boldsymbol{F}$ a tensor of impulsive forcing terms, directed like the reference axes, and L an operator that combines Newton's law and Hooke's law (Snieder 1986a, or Snieder 1987, chapter 2, eqs 15 and 16). First, a JWKB solution of eq. (1) is found (in the far-field approximation); let us call it $\boldsymbol{G}_0$; we can think of it as the zeroth-order solution.

We next use $\boldsymbol{G}_0$ to find the effects $\delta\boldsymbol{G}$ on the Green tensor, of small 3-D perturbations in the Earth's elastic parameters [density ρ (x), seismic velocities v_P(x) and v_S(x), with x denoting position]. We replace ρ, v_P and v_S with $\rho + \delta\rho$(x), $v_P + \delta v_P$(x) and $v_S + \delta v_S$(x) in the formula for L, thus obtaining the perturbed operator $L + \delta L$. The perturbed problem is then

$$(L + \delta L)(\boldsymbol{G}_0 + \delta\boldsymbol{G}) = \boldsymbol{F}. \tag{2}$$

After subtracting eq. (1) from (2), and neglecting second-order terms,

$$L\delta\boldsymbol{G} + \delta L\boldsymbol{G}_0 = \boldsymbol{0}, \tag{3}$$

or

$$L\delta\boldsymbol{G} = -\delta L\boldsymbol{G}_0, \tag{4}$$

which is equivalent to the original Green's problem (1), with $\boldsymbol{F}$ replaced by the non-impulsive forcing $-\delta L\boldsymbol{G}_0$. By virtue of the properties of the Green tensor, $\delta\boldsymbol{G}$ is then found by a simple convolution of $\boldsymbol{G}_0$ with the new forcing term $-\delta L\boldsymbol{G}_0$ (Snieder 1986a, 1987, chapter 2, eqs 17–21).

A similar approach is followed by Dahlen *et al.* (2000), and implicitly by Montelli *et al.* (2004a,b). Its major shortcoming is that, as Dahlen *et al.* (2000) themselves noted, the solution is 'grounded upon an approximate JWKB solution': the infinite-frequency limit is not explicitly removed, but dealt with by perturbing linearly to first-order the zeroth-order solution found in that limit; there is no proof that this should accommodate the non-linear relation between seismic measurements and the Earth's seismic properties, that exists in a finite-frequency regime (i.e. in the real world). Hung *et al.* (2000), however, have validated the results of Dahlen *et al.*'s (2000) scattering theory by comparison with numerical calculations, where no approximation is required.

After some algebra, Snieder (1986a) found a relation (involving the summation of normal modes before and after scattering) between $\delta\boldsymbol{G}$ and $\delta\rho$(x), δv_P(x) and δv_S(x), which forms the basis of a tomographic inverse problem: a seismic measurement can be expressed in terms of $\delta\boldsymbol{G}$, while δv_P, etc., form the unknown tomographic model to be derived (again, the reasoning of Dahlen *et al.* 2000, is qualitatively analogous).

A further step is required if maps of surface wave phase velocity (rather than δv_P, etc.) are to be inverted for. Working in the far-field approximation (Snieder 1986b, 1987, chapter 3, eqs 8.2 and 8.3) shows that conversions between different modes can be neglected if seismic heterogeneities are smooth and relatively weak; he can then relate directly $\delta\boldsymbol{G}$, and therefore a surface wave measurement, to the phase velocity associated with a certain surface wave mode (Snieder 1986b, or 1987, chapter 3, eq. 9.3; Snieder 1987, chapter 8, sections 2 and 3; Snieder 1988; Zhou *et al.* 2004, Section 6 and eq. 6.2). This leads to eq. (9) of Spetzler *et al.* (2002), who used Snieder's (1987) treatment to introduce phase-velocity Fréchet kernels, that is, a function $K(\omega, \theta, \phi)$ such that a relative phase-anomaly measurement $\delta\varphi/\varphi(\omega)$ can be written as

$$\frac{\delta\varphi(\omega)}{\varphi(\omega)} = \int_\Omega K(\omega, \theta, \phi)\frac{\delta c(\omega, \theta, \phi)}{c(\omega)}\, d\Omega, \tag{5}$$

where θ and ϕ denote colatitude and longitude, respectively, ω frequency, Ω the unit sphere and c phase velocity (which here I assume to be laterally constant in the reference model). The Fréchet kernels $K(\omega, \theta, \phi)$ are also often referred to as sensitivity functions, partial derivatives, 'banana-doughnuts', etc. Based on their eq. (9), Spetzler *et al.* (2002, eq. 16) found algebraically an analytical expression for $K(\omega, \theta, \phi)$, namely

$$K(\omega, \theta, \phi) = -\frac{\sin(\theta)\sqrt{R\sin(\Delta)}}{\Delta}\frac{\sqrt{\omega}\sin\left[\frac{\pi\omega R}{c(\omega)}(\theta - \frac{\pi}{2})^2\frac{\sin(\Delta)}{\sin(\phi)\sin(\Delta-\phi)} + \frac{\pi}{4}\right]}{\sqrt{c(\omega)\sin(\phi)\sin(\Delta-\phi)}}, \tag{6}$$

valid in a reference frame where both source and station lie on the equator, and with Δ denoting epicentral distance and R the Earth's radius.

Like Snieder (1987), Zhou *et al.* (2004) proceeded by mode summation in the far-field approximation; they, however, account for coupling between modes, to derive sensitivity functions relating surface wave measurements to 3-D heterogeneities in shear and compressional velocity, and density. Only in a subsequent step (section 6 of their paper), they started to neglect mode coupling: this allowed them to reduce their 3-D sensitivity functions to 2-D, and establish a relationship between surface wave measurements and

2-D heterogeneities in phase velocity; the results of their procedure are consistent with those of Spetzler *et al.* (2002).

In the following, I will use Spetzler *et al.*'s (2002) formula for $K(\omega, \theta, \phi)$, projecting it on a pixelized, rather than harmonic parametrization.

3 BORN-THEORY TOMOGRAPHY

I implement eq. (6) and repeat Spetzler *et al.*'s (2002) calculation of scattering-based Fréchet kernels at a discrete set of teleseismic epicentral distances (20°–179°, with 1° increments), for a source located at (0°, 0°) and stations along the equator. Examples are given in Fig. 1, for two different periods and two epicentral distances. Having been derived in the far-field approximation, expression (6) for $K(\omega, \theta, \phi)$ is singular near source and receiver, which may lead to major problems in tomographic applications; the singularity is taken care of, somewhat arbitrarily but effectively, by replacing scatterer-to-source and scatterer-to-receiver distances with some small, but finite, ε as they approach 0 (J. Spetzler, private communication, 2004). Also, following Spetzler *et al.* (2002) I average the Fréchet derivatives over a frequency band of $2\Delta\omega = 5$ mHz around the chosen frequency ω, 'since phase-velocity measurements at a single frequency are not possible owing to the finite sampling of the seismograms and the finite parametrization of the dispersion curve in the measurement process' (Spetzler *et al.* 2002). The chosen value for $2\Delta\omega$ approximately equals the accuracy of dispersion measurements considered here (i.e. the spacing between splines parametrizing the measured dispersion curves in Ekström *et al.* 1997, fig. 1). Incidentally, averaging over frequency has a smoothing effect on $K(\omega, \theta, \phi)$, damping its oscillations away from its first maximum.

I have tested my code against that written by Spetzler *et al.* (2002), and found coincident results.

Confirming previous studies on the subject, scattering kernels are maximum *away* from the JWKB ray path (which, in the case of constant reference phase velocity, coincides with the great-circle arc connecting source and station), and small (but non-zero) on the ray path itself (Fig. 1). The reader is referred to Spetzler *et al.* (2002, section 2.4, fig. 2) for a more thorough discussion of Fréchet kernels and their properties.

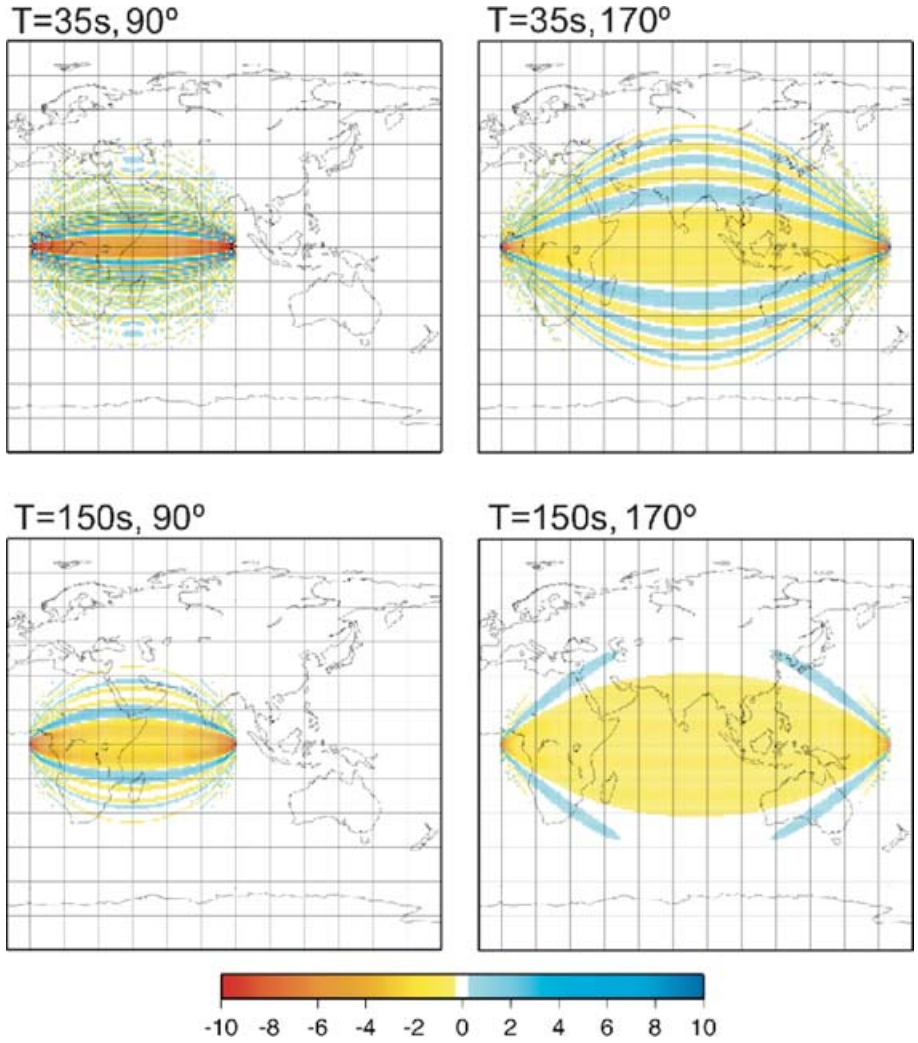

Figure 1. Fréchet kernels, for epicentral distances $\Delta = 90°$ and 170°, and surface wave periods $T = 35$ s and 150 s (as indicated), from Spetzler *et al.*'s (2002) expression (coincident for Love or Rayleigh waves). At any location, the plotted value equals the ratio of a relative phase anomaly observed at the station, to a localized relative phase-velocity anomaly at that location, for that source–station geometry and in the absence of other heterogeneities. I first compute kernels for one source at (0°, 0°) and equally spaced stations along the equator, then interpolate and rotate them to account for real source–station geometries (see Fig. 2).

After finding Fréchet kernels and storing them on disk, the next step is to set up the tomographic inverse problem

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{d}, \tag{7}$$

where x is the vector of solution coefficients (each of its entries the constant value of $\delta c/c$, at the chosen frequency, within a certain pixel), d the vector of relative phase anomalies, and

$$A_{ij} = \int_{j\text{th pixel}} K_{i\text{th datum}}(\omega, \theta, \phi)\, d\Omega \tag{8}$$

(Woodhouse & Girnius 1982, eqs 56 and 57).

I have then written an algorithm that (i) reads phase anomalies from Ekström *et al.*'s (1997) database, updated and averaged into 'summary' data as mentioned by Boschi & Ekström (2002), (ii) finds for each datum the appropriate $K(\omega, \theta, \phi)$ by spline interpolation between Fréchet kernels previously calculated at discrete epicentral distances, (iii) rotates $K(\omega, \theta, \phi)$ from the equator to the source–station geometry associated with the datum in question, and projects it onto the grid of pixels defined above (Fig. 2), to find **A** and $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$.

Rotating pre-computed kernels is faster than computing a kernel directly for each source–station couple in the database.

The matrix $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$ is useful for a rigorous evaluation of model resolution (e.g. Boschi 2003), and in case a direct, rather than iterative inversion algorithm is to be applied to find the least-squares solution of (7). Here, I have experimented with both LSQR (iterative) and Cholesky factorization (direct), to find coincident results; because of the sparsity of **A** (denser than in the ray-theory case, but still very sparse), LSQR is much faster and thus preferable.

4 REGULARIZATION AND TRADE-OFF

I next invert the database of Ekström *et al.* (1997), updated as explained by Boschi & Ekström (2002), to derive phase-velocity maps both in the Born-theory formulation described above, and the traditional JWKB approximation (e.g. Ekström *et al.* 1997).

In principle, if database, parametrization, units of measurement and regularization are the same in the two cases, any discrepancy between the maps can only result from the different order of approximation applied. It is not clear, however, what it means to *equally* regularize the Born- and ray-theory inverse problems. One might think of scaling the damping parameters based on the average size of the diagonal entries of $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$, but a closer look at $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$ discourages it: Fig. 3 shows that the introduction of Born theoretical kernels tends to lower the mean value of diagonal entries, but simultaneously broadens their range, with higher maxima and lower minima. This is the expected effect of a redistribution of tomographic resolution, owing to the higher sensitivity in the immediate vicinity

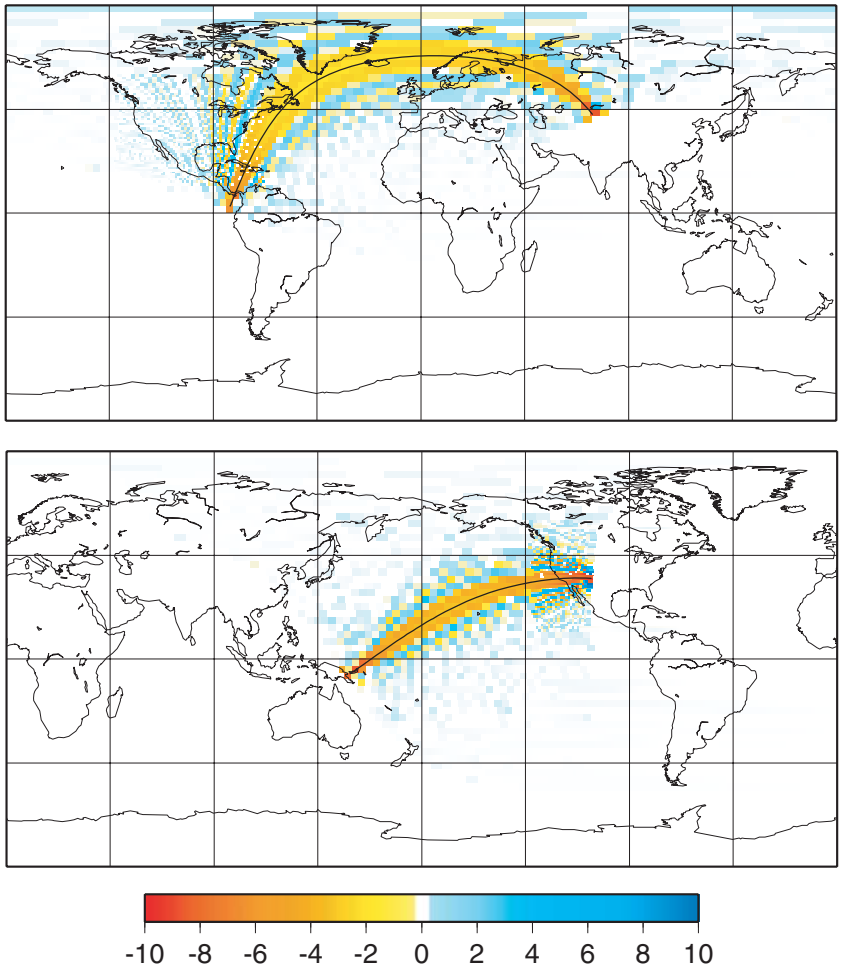

Figure 2. Two examples of Fréchet kernels rotated and projected onto my pixel grid. The oscillatory behaviour of kernels is more faithfully reproduced within the more finely parametrized region (North America). Shown here are the average values of kernels within each pixel; this is different from the corresponding entries of A, which involve multiplication by the pixel area.

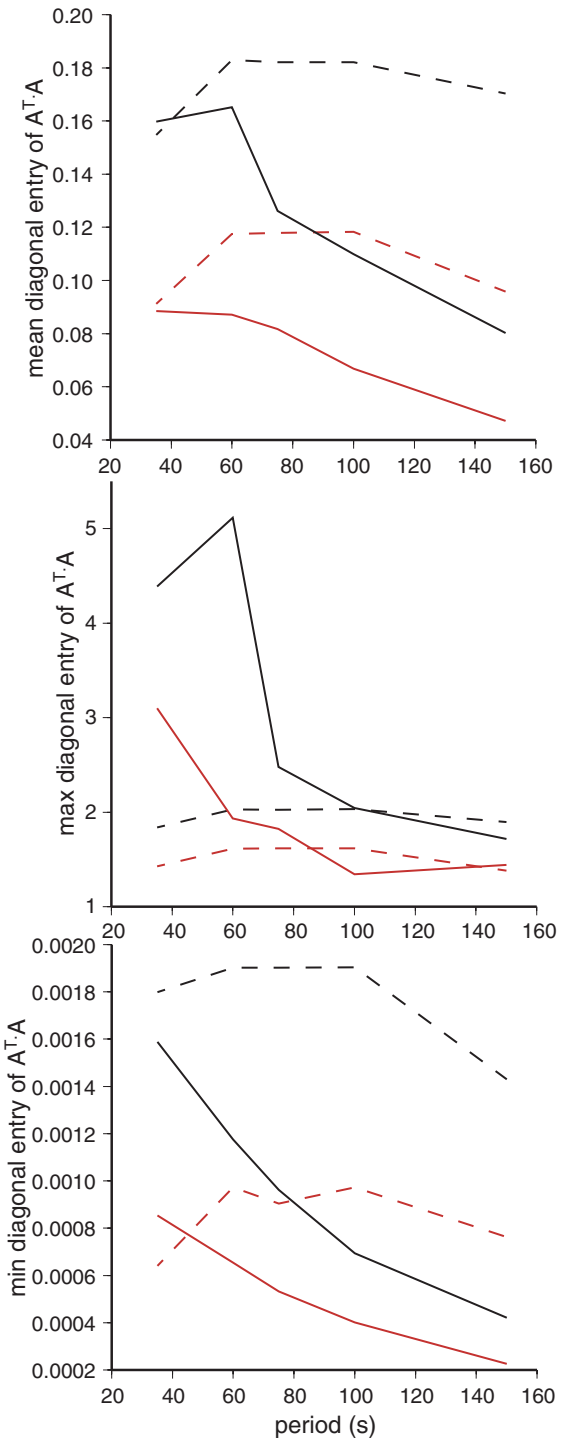

Figure 3. Values of the mean (top panel), largest (middle panel) and smallest (bottom panel) diagonal entries of $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$, as a function of period for Rayleigh (black lines) and Love (red) waves, in the finite-frequency (solid) and JWKB (dashed) cases.

of sources and stations, where $K(\theta, \phi)$ is singular. Particularly at shorter periods, a lower damping in the Born-theory case, while justified by the smaller average of $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A}$ diagonal entries, would then result in the solution being locally underdamped because of their higher maxima. Other comparative studies of tomographic methods (e.g. Spetzler *et al.* 2001, 2002; Montelli *et al.* 2004b) are bound to encounter the same difficulty.

In many applications of inverse theory, 'trade-off curves' are employed to establish an acceptable range of regularization schemes. After performing numerous differently regularized inversions, a measure of the misfit of each solution model to the data (variance, χ^2, ...) is plotted against a measure of the complexity of the model itself (model RMS, integrated norm of its gradient, ...). If a sufficiently broad spectrum of regularization schemes has been spanned, the resulting curve should vaguely resemble the letter 'L' (see, e.g. Ekström *et al.* 1997, fig. 11), and is therefore sometimes dubbed 'L-curve' (Hansen 1992): in a regime of exceedingly strong regularization, a small growth in model complexity is typically sufficient to reduce significantly the misfit, while if the regularization constraint is too weak, the misfit remains approximately constant even after a large growth in model complexity; the solution being dominated by the effects of data noise and numerical instabilities. The preferred solution should be chosen near the corner of the trade-off curve.

Trade-off curves are by definition independent of damping; rather, they describe the dependence on damping, of the least-squares solution to the inverse problem in question. If one finds L-curves separately for both the Born-theory and ray-theory inverse problems associated to the same database, it is then legitimate to compare the two curves.

I have derived trade-off curves for the surface wave phase-velocity inverse problems formulated here, prescribing no norm minimization constraint and gradually decreasing the roughness minimization parameter. Arguments for preferring roughness over norm minimization are given, for example, by Inoue *et al.* (1990) or Boschi & Dziewonski (1999). The roughness damping matrix is defined as in Boschi & Dziewonski (1999), but accounting for the non-uniformity of the grid and inherent changes in pixel area. Sample trade-off curves from Born and ray theoretical inversions at four selected surface wave modes are shown in Fig. 4. If we neglect phase-velocity maps with gradients too high to be physical, Fig. 4 shows that Born-theory solutions systematically achieve a worse datafit than ray-theory solutions of equal complexity. This result is, at least, counterintuitive, as one would expect a better theoretical formulation of the inverse problem to provide a better solution model, at any level of model complexity. It is, nevertheless, consistent with the trade-off analysis of Montelli *et al.* (2004b, fig. 7), and some of the 2-D tomography results of Zhou *et al.* (2005, fig. 20). Boschi *et al.* (2006) attempted an explanation in terms of the Akaike information criterion (Akaike 1974), and suggested that the higher complexity (at any given level of datafit) of Born-theory tomographic images might reflect true complexities in the Earth's structure, not resolved by simple JWKB theory.

4.1 Model selection by L-curve analysis

A visual analysis of L-curves is not sufficient to identify their corner, whose location depends on the definition of image complexity and on the exaggeration of the axes. Apparently small changes in complexity and misfit can result in very large perturbations, particularly to the short spatial wavelength component of solution models. Following Zhou *et al.* (2005), I define image complexity as the ratio of integrated model roughness to model norm. After this normalization, I calculate the curvature (e.g. Weisstein 2003) of each L-curve, and select the solution model of complexity and misfit corresponding to maximum curvature. Fig. 5 shows that points of maximum curvature are systematically well defined and easily identified.

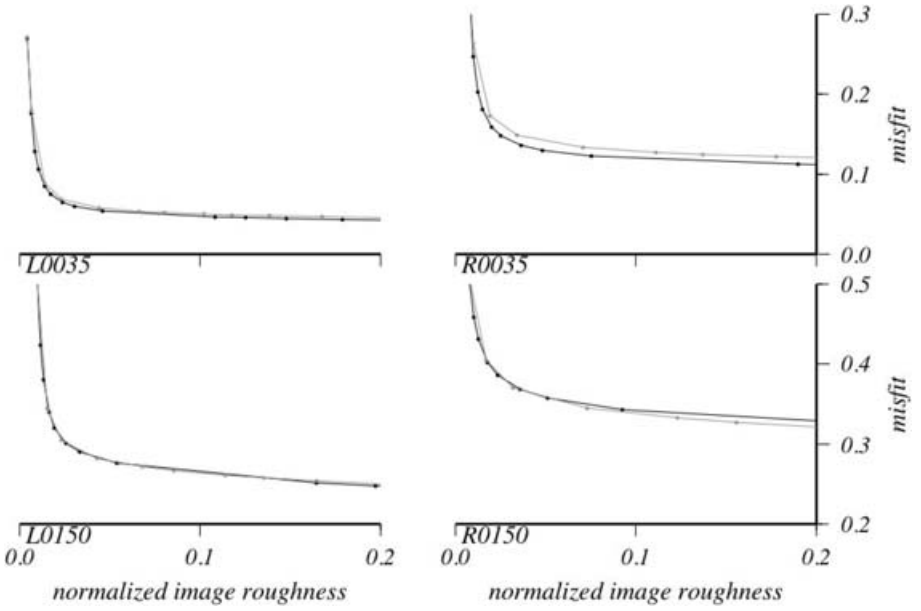

Figure 4. Trade-off curves (L-curves) resulting from ray- (black lines) and Born-theory (grey) inversions of the Harvard dispersion database: Love (left-hand panel) and Rayleigh waves (right-hand panel) at 35 s (top panels) and 150 s (bottom panels) periods, respectively. The image roughness is defined as the squared modulus of the dot-product of roughness damping matrix times vector of model coefficients. I normalize it, dividing by the total sum of squared model coefficients. Misfit equals 1 - variance reduction.

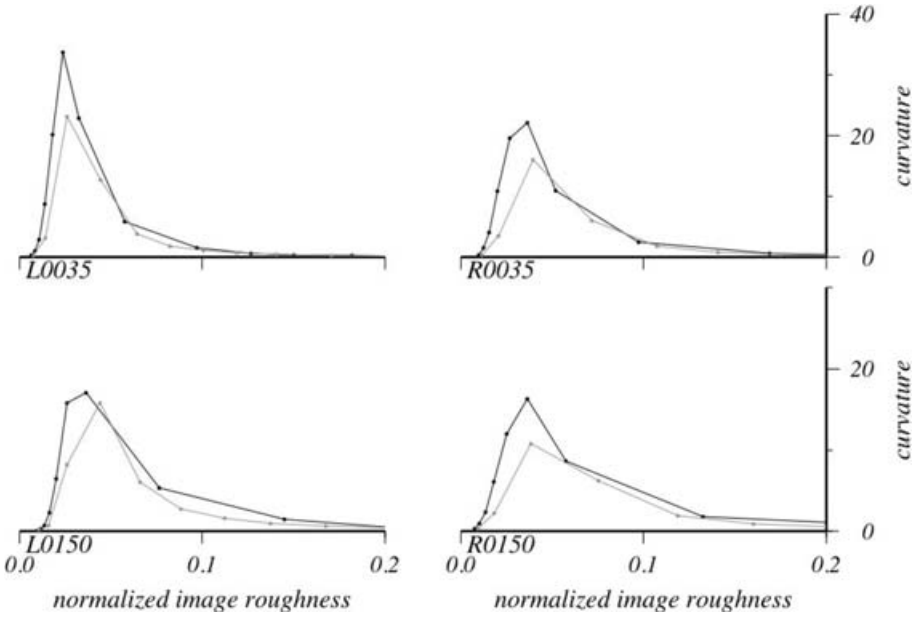

Figure 5. Curvature of the L-curves from Fig. 4. For each L-curve, the point of maximum curvature identifies my preferred solution.

I apply the same criterion to find optimal Born and ray theoretical solutions: the resulting maps can be considered almost equivalently regularized, and it is then legitimate to compare them visually, to evaluate quantitatively the effects of scattering on tomography.

4.2 Synthetic test

Substituting in eq. (5) a realistic *a priori* model $\delta c/c$, and $K(\omega, \theta, \phi)$ as defined by (6), I compute at each considered surface wave frequency a set of theoretical phase anomalies, with the same source–station geometry as the Harvard database. I use a Box–Muller transformation (e.g. Weisstein 2003) to randomly generate sets of normally distributed measurement errors, with standard deviation equal to the estimated, real observational uncertainty at the corresponding surface wave frequency (Ekström *et al.* 1997, table 2). After adding such synthetic errors to all synthetic phase-anomaly observations, I invert them with the same algorithm applied to real data, and conduct on the resulting output models a trade-off analysis as above; the corresponding L-curves, and their curvatures, are shown in Figs 6 and 7, respectively; they indicate that inverting phase-velocity data through Born rather than ray theory should result in a visible vertical shift of the L-curves, with Born-theory solutions achieving a better datafit than ray-theory ones at any level of model complexity.

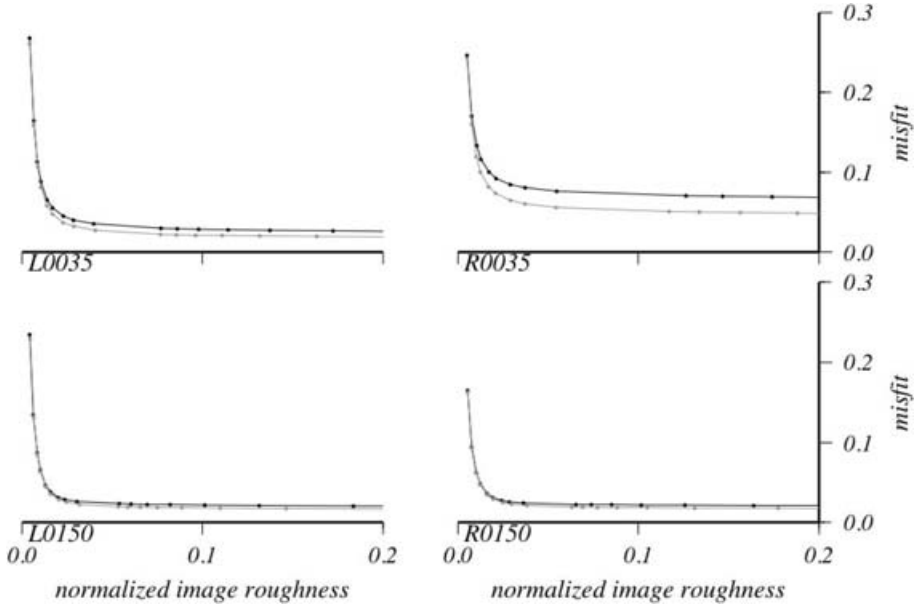

Figure 6. Same as Fig. 4, but solutions are derived from the inversion of *synthetic* data, as described in Section 4.2.

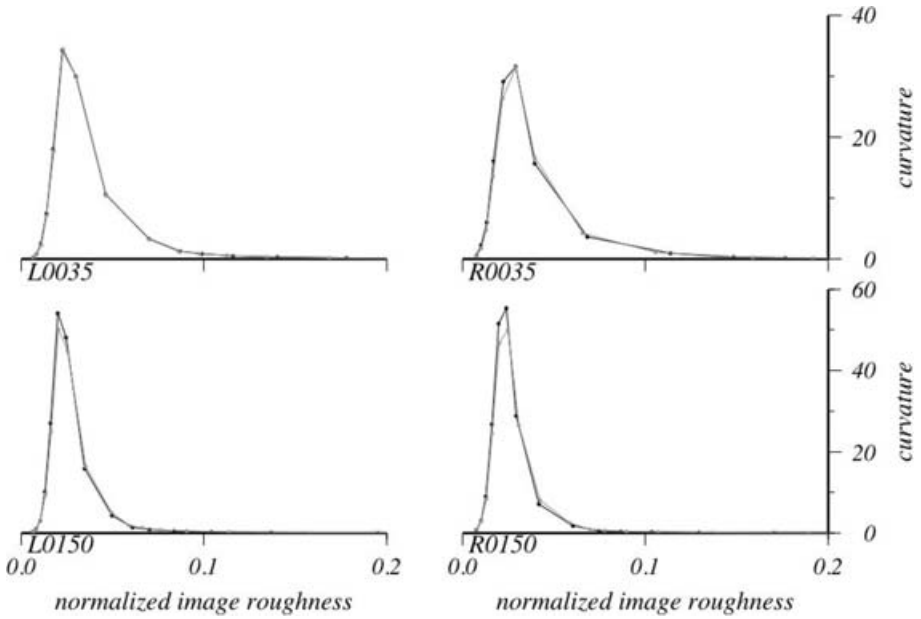

Figure 7. Curvature of the L-curves from Fig. 6.

Figs 6 and 7 are not consistent with Figs 4 and 5, where the said vertical shift is either reversed in sign, or only episodical and barely visible. I must infer that, at least at the global scale, the quality of phase-velocity models is most likely *not* improved by the introduction of a Born theoretical formulation. This unexpected result could be explained considering that comparing Figs 4–6, or Figs 5–7, is legitimate only in the assumption that the Born theoretical formulation be *exact*. This assumption is only valid within a certain approximation, because (i) phase velocity can only be defined for a relatively smooth Earth, and in the neglect of normal-mode coupling, with the implications well pointed out by Zhou *et al.* (2005, section 5); (ii) here, and in most other published Born theoretical tomographic studies (e.g. Li & Romanowicz 1996; Spetzler *et al.* 2002; Montelli *et al.* 2004a,b; Zhou *et al.* 2005), sensitivity kernels are derived perturbing approximate, far-field JWKB solutions of the Earth's momentum equation, and are subsequently not strictly correct in the vicinity of sources and stations (Favier *et al.* 2004), where they become singular (I address this issue in Appendix B below); (iii) surface wave energy propagates, to a large extent, within the Earth's crust, which we know to be very heterogeneous both laterally and vertically: surface wave propagation must then be a non-linear phenomenon, while Born-theory tomography still relies on a linearization of the inverse problem.

Alternatively, or additionally, the improvement in model quality achieved by trading a ray theoretical for a Born theoretical formulation is limited by parametrization and data coverage. Fig. 2 shows that my pixel grid describes $K(\omega, \theta, \phi)$ fully well only within the high-resolution region, while datafit and image complexity

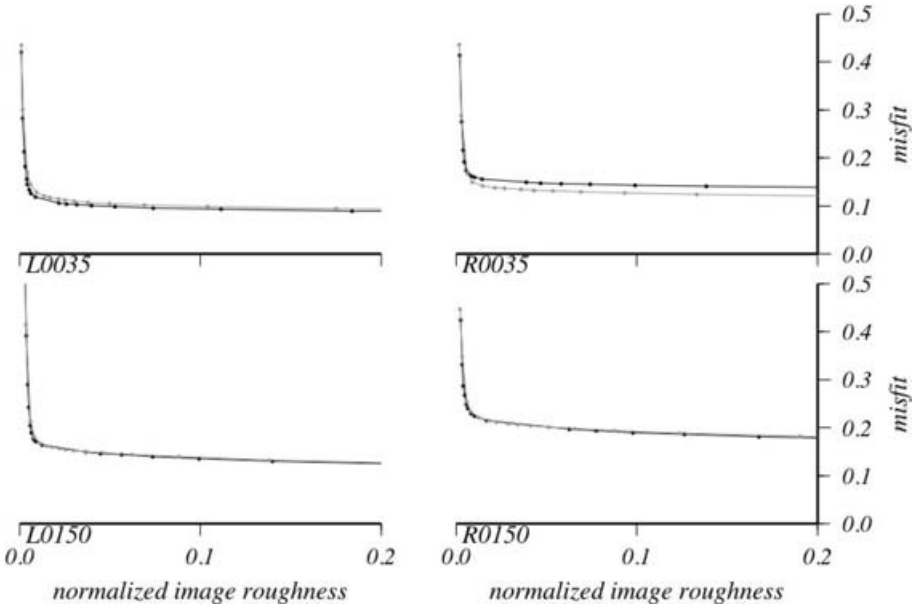

Figure 8. Same as Fig. 4, but solutions are derived from the inversion of data whose JWKB ray path samples the high-resolution region, only. Accordingly, only model coefficients associated with the high-resolution region are considered in the calculation of model roughness (Section 4.3).

in Figs 4–7 are calculated globally. Within North America, on the other hand, the geographic coverage of the database I invert, while very good, might still be inadequate to constrain short-spatial-wavelength heterogeneities, making scattering kernels practically ineffective.

4.3 L-curve analysis, high-resolution region

I extract from the inverted database phase anomalies associated with JWKB ray paths sampling the more finely parametrized region (on average, roughly a quarter of the database). I carry out a new trade-off analysis, plotting, for each solution model, its cumulative misfit to those data only, versus its roughness integrated only over the high-resolution region itself. The results are shown in Figs 8 and 9. In comparison with Fig. 4, both Born and ray theoretical L-curves are shifted significantly downwards with respect to the ray theoretical ones, indicating that refining the parametrization helps improving the datafit without necessarily increasing image complexity.

Born and ray theoretical L-curves are, again, approximately coincident, except for the case of short- to intermediate-period Rayleigh waves (only the 35-s case is shown in Figs 8 and 9 as an example), where the Born L-curve is well below the ray-theory one.

In summary, resolution limits enumerated in the previous section are at least partially overcome by a refinement of nominal resolution.

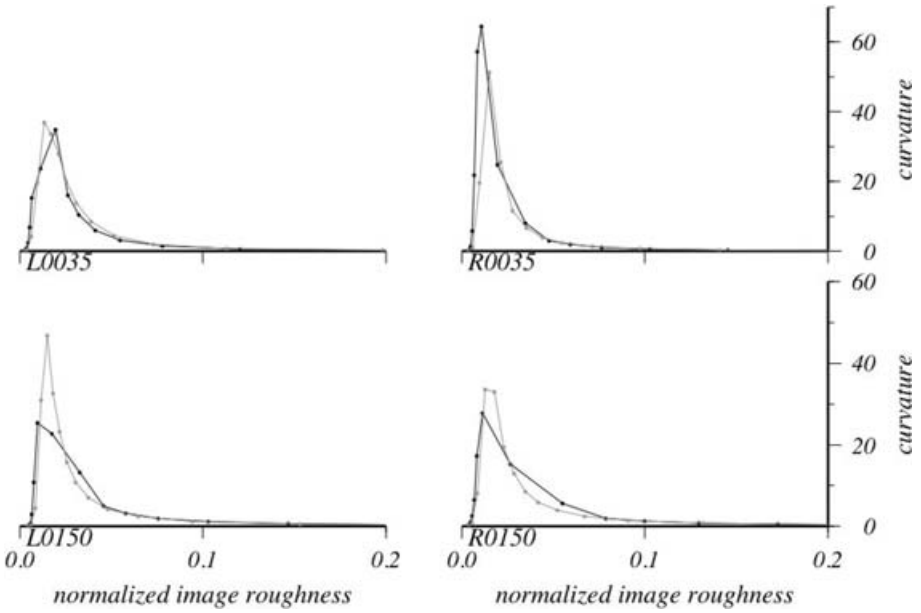

Figure 9. Curvature of the L-curves from Fig. 8. Solutions corresponding to maximum curvature are the same, or approximately the same as those found from Fig. 5.

5 RESOLUTION MATRICES

Tomographic resolution is generally limited by data quality and coverage, adequacy of the parametrization, and accuracy of the theoretical formulation used to relate seismic measurements to the Earth structure. The conservative data selection criteria of Ekström *et al.* (1997) guarantee that data quality is high (Carannante & Boschi 2006), particularly at relatively long periods (100 s and more), where 'phase measurements. . . can be associated with a total cycle count without difficulty' (Ekström *et al.* 1997). I have intentionally overparametrized the solution, and proved (Fig. 2, Appendix A) that, at least in North America, pixels are sufficiently small to reproduce the fast oscillations of sensitivity functions. It remains to be verified that the data coverage is sufficiently dense and uniform.

Denoting $\mathbf{D}$ as the roughness damping matrix, and λ the selected damping parameter, I find the model resolution matrix

$$\mathbf{R} = (\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A} + \lambda \mathbf{D})^{-1} \cdot \mathbf{A}^{\mathrm{T}} \cdot \mathbf{A} \tag{9}$$

after a Cholesky factorization of $\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A} + \lambda \mathbf{D}$ (e.g. Boschi 2003), both in the Born- and ray-theory cases. Independent values of λ for the Born and ray theoretical inversions are determined from Fig. 5 as explained above.

Resolution matrices associated with equivalently regularized, Born- and ray-theory inversions of 35-s Love wave measurements are shown, as an example, in Fig. 10; since differences in R in the two cases are very subtle, I prefer to focus on two chosen rows, rather than plotting the entire matrix like Boschi (2003). The *j*th entry of the *i*th row of $\mathbf{R}$, R_{ij}, is a measure of the fictitious trade-off ('smearing') between the *i*th and *j*th model coefficients (Boschi 2003; Soldati & Boschi 2005), and plotting, for all values of j, R_{ij} on the *j*-th pixel, I obtain a more easily readable image of fictitious coupling between model coefficients.

In the case of perfect resolution, maps in Fig. 10 should be entirely white, except for one red pixel at $R_{ii} = 1$. For pixels within the high-resolution region (top and middle panels of Fig. 10), typically $R_{ii} \ll 1$, and $R_{ij} \sim 1$ for numerous pixels in the vicinity of the *i*th one (located 2° South of Lake Michigan). This means that, as was to be expected, heterogenities of lateral extent comparable to the gridsize cannot be properly mapped, regardless of the theoretical approach. Larger heterogeneities, however, can be imaged without a strong loss in amplitude resolution (I find, in fact, that the size of heterogeneities mapped within the region of interest is comparable to those outside, as in Boschi *et al.* 2006), as $R_{ij} > 0$ and neighbouring pixels can coalesce to form a coherent anomaly of longer spatial wavelength.

Outside the region of interest, in areas relatively well covered by the data $R_{ii} \sim 1$ and $R_{ij} \ll 1$, smearing is, however, distributed over a larger number of pixels (bottom panels of Fig. 10), which confirms the utility of a locally finer parametrization.

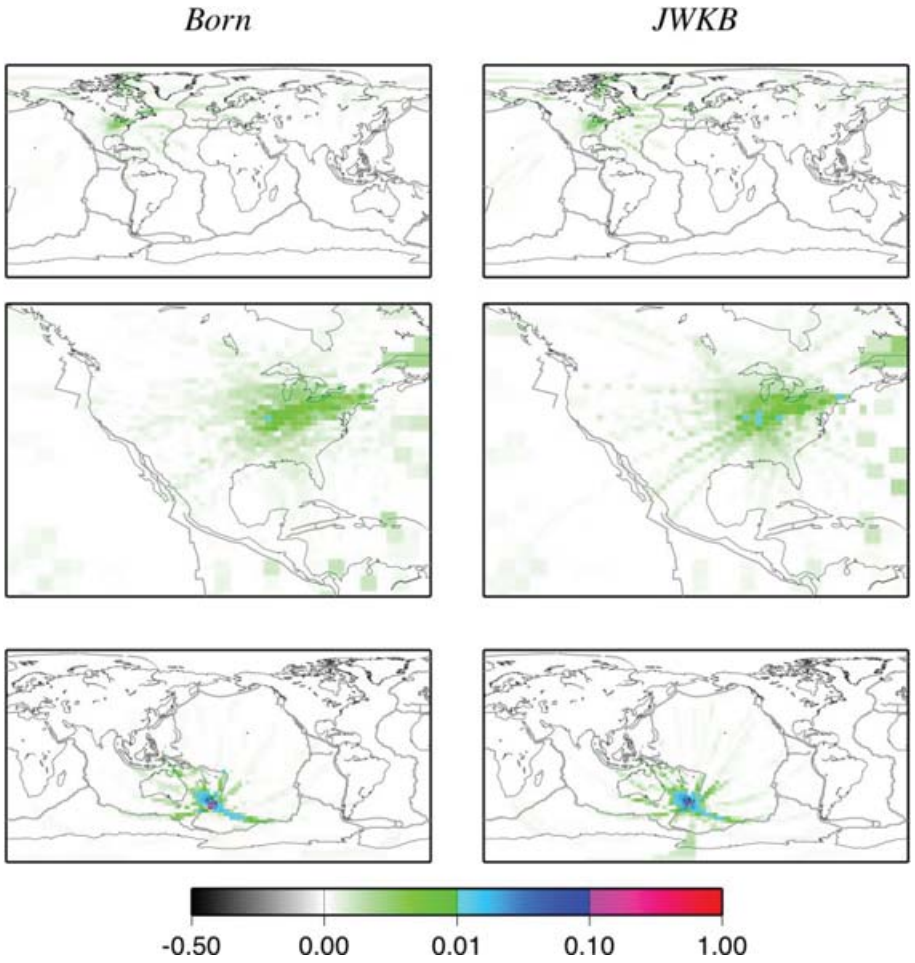

Figure 10. Top panels: row of the 35-s Love wave resolution matrices R associated with a pixel located in northern-central America; middle panels: detail of top panels; bottom panels: row of R associated with a pixel in New Zealand. Resolution matrices found in the scattering- (left-hand panels) versus ray-theory (right-hand panels) approximations are compared. For each value of *j*, the colour of the *j*th pixel depends on the value of R_{ij}; R_{ij} is a measure of fictitious trade-off between *i*th and *j*th model parameters.

While coupling between close pixels follows a slightly different geometry (with the expression of ray paths clearly visible in the ray-theory example), the total amount of trade-off in Born- versus ray-theory inversions appears to be practically the same.

It would be possible to enhance **R** both globally and locally, by relaxing regularization, and still find 'acceptable' models; the relatively low resolution suggested by Fig. 10 is an expression of my conservative choice of regularization, according to the criterion set forth in Section 4.1.

6 PHASE-VELOCITY MAPS

The finer parametrization adopted in North America results in a downward shift of L-curves associated with data sampling the high-resolution region, with respect to those derived from the entire data set (Section 4.3); this suggests that the growth in nominal resolution has effectively improved model quality.

A visual inspection of optimally regularized phase-velocity maps in Figs 11 (Love waves) and 12 (Rayleigh), selected as explained in Section 4, confirms that the fine, $1^\circ \times 1^\circ$ grid helps resolving expected, large velocity gradients, that a coarser parametrization (Boschi & Ekström 2002; Boschi *et al.* 2006) would not have reproduced so faithfully.

Of the surface wave modes I have considered, Love waves at 35 s are most sensitive to the properties of the crust (Boschi & Ekström 2002). The corresponding ray- and Born-theory maps (Fig. 11) are both characterized by a strong contrast between slow continental crust, and fast young oceanic crust, that matches geophysical expectations; the boundary of the imaged slow continental heterogeneity follows closely the Western North America coastline.

At 150 s period, Love (as well as Rayleigh) wave propagation is more importantly affected by lateral heterogeneities in the Earth's lithosphere and upper mantle. Accordingly, phase-velocity anomalies found at this period from both approaches can be correlated with free-air gravity observations (Nerem *et al.* 1994, 1995; Simons & Hager 1997; Perry *et al.* 2003), or with the distribution of surface heat flow (Nataf & Ricard 1996, fig. 7; Godey 2002, fig. 5.3).

Last, the Rayleigh wave maps of Fig. 12 are consistent with Nataf & Ricard's (1996, fig. 5) regionalization of continents, with a slow region corresponding to 'tectonic continent', a fast craton, and $\delta c \sim 0$ within the 'stable platform' in south-central and Eastern America (compare with fig. 5.4 of Godey 2002). Rayleigh wave 150-s maps (bottom of Fig. 12) reproduce particularly well the boundary between craton and platform along the Ontario and Erie lakes.

While it is encouraging that tomographic images reflect independent geophysical observations at the regional scale, there is no evidence here that the introduction of Born-theory results in a significant improvement with respect to the simpler, ray theoretical method: all the mentioned expected features are resolved equally well by both algorithms. While a small difference exists in the

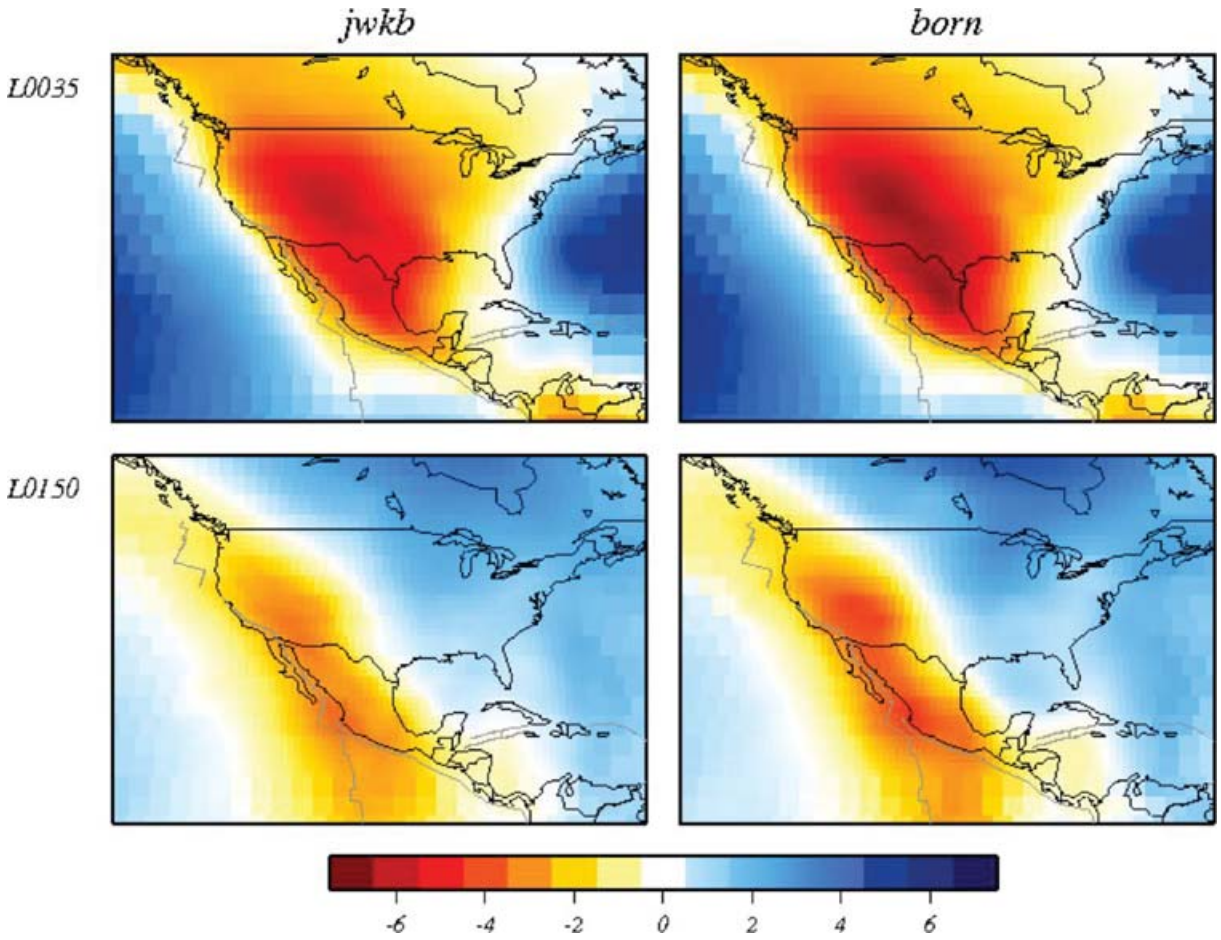

Figure 11. Phase velocity (expressed as per cent perturbation to PREM) of 35-s (top panel) and 150-s (bottom panel) Love waves, from the ray- (left-hand panel) and scattering-theory (right-hand panel) formulations.

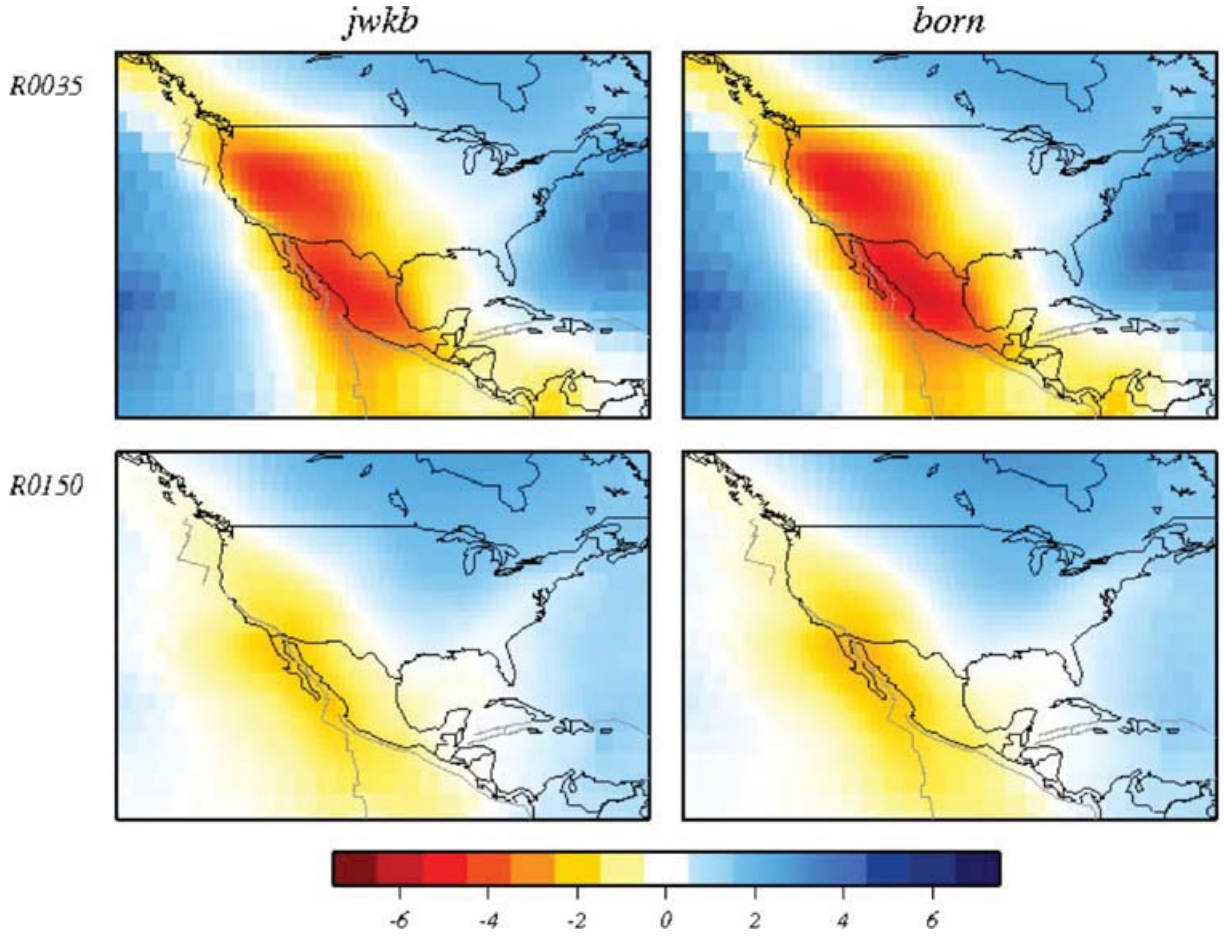

Figure 12. Same as Fig. 11, Rayleigh waves.

amplitudes of heterogeneities mapped by Born- versus ray-theory tomography (the former being systematically larger), the heterogeneity patterns derived from the two approaches can be considered coincident.

7 CONCLUSIONS

Following the works of Snieder (1987) and Spetzler *et al.* (2002), I have made use of scattering (Born, banana-doughnut,. . .) theory to derive global Love and Rayleigh wave phase-velocity maps from the updated Harvard fundamental-mode dispersion database (Ekström *et al.* 1997).

To investigate the improvement in model quality achieved, thanks to Born theory, I compare Born and ray theoretical tomographic images, derived from the same database and with the same parametrization and inversion algorithm. Born theory should enhance tomographic resolution, to correctly map anomalies of extent smaller than the wavelength of inverted observations (Spetzler *et al.* 2001, 2002; Sieminski *et al.* 2004); for a Born- versus ray-theory comparison to be meaningful, it is then crucial that the parametrization be fine enough to resolve such anomalies (Fig. 13, Appendix A). I make use of a non-uniform pixel grid, finer over North America, a region particularly well covered by the data.

To correctly evaluate the significance of differences between Born and ray theoretical tomographic results, one must make sure that the procedures devised in the two cases are otherwise exactly equivalent: parametrization, regularization and inversion algorithms should be the same. It is easy to write Born and ray theoretical tomographic software based on the same pixel grid and, for example, LSQR routine; but regularization is a trickier issue. Different sensitivity functions redistribute element amplitudes in the matrix to be inverted, and regularization must be redefined accordingly. I show in Fig. 3 that this cannot be reduced to a simple scaling of the damping parameters, and propose in Section 4.1 a criterion to identify equivalently regularized Born and ray theoretical solutions, that can be compared. My approach rests on the trade-off, or L-curve (Hansen 1992) analysis, used to determine acceptable regularization schemes in damped inverse problems. In practice, I plot the fit achieved by solution models against a measure of their complexity, and select as optimal model the one corresponding to the point of maximum curvature of the plot. At any given surface wave mode, ray and Born theoretical phase-velocity maps selected with this criterion can be considered equivalently regularized and can be compared.

The comparison between Born- and ray-theory trade-off curves (Fig. 4) shows that images based on Born theory are systematically more complicated than ray-theory ones, explaining the data equally well; or, a Born-theory solution systematically reduces data variance less than the ray-theory solution of equal complexity (see also Boschi *et al.* 2006). This result contradicts theoretical expectations based on a synthetic test, and in the assumption that the Born-theory solution be close to exact (Section 4.2). This situation

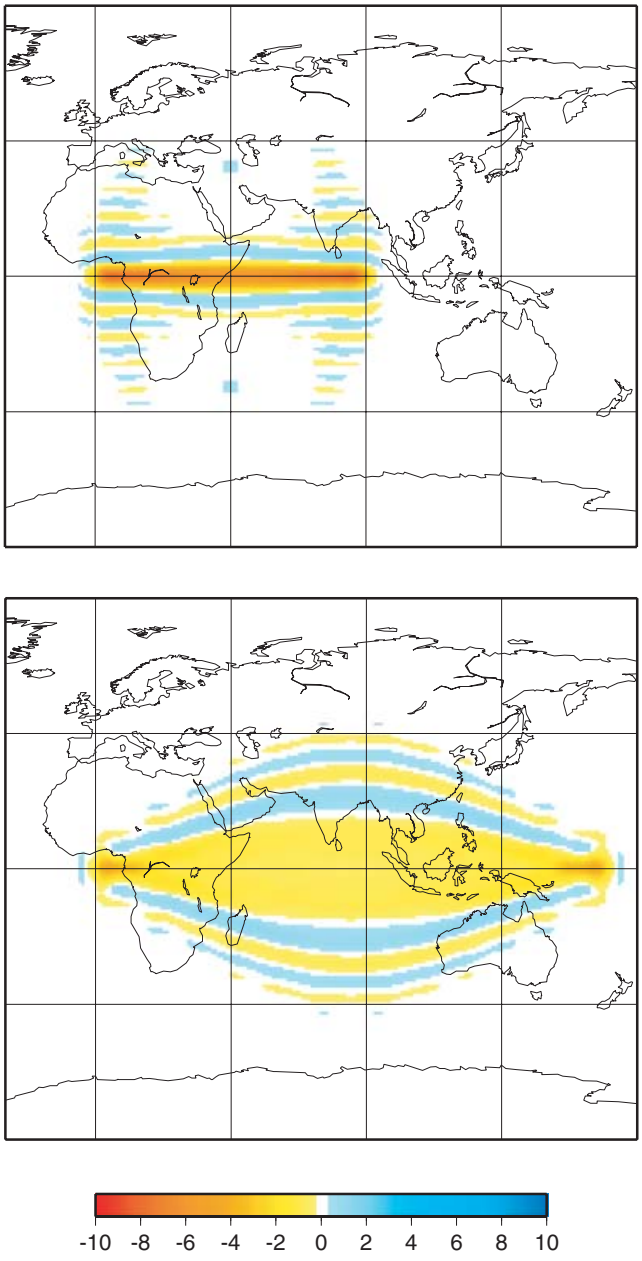

Figure 13. Spherical harmonic expansions, up to degree 40, of 35-s Fréchet kernels, at epicentral distances 90° (top panel) and 170° (bottom panel). Compare with the top panels of Fig. 1: the harmonic expansion tends to smooth the kernels significantly.

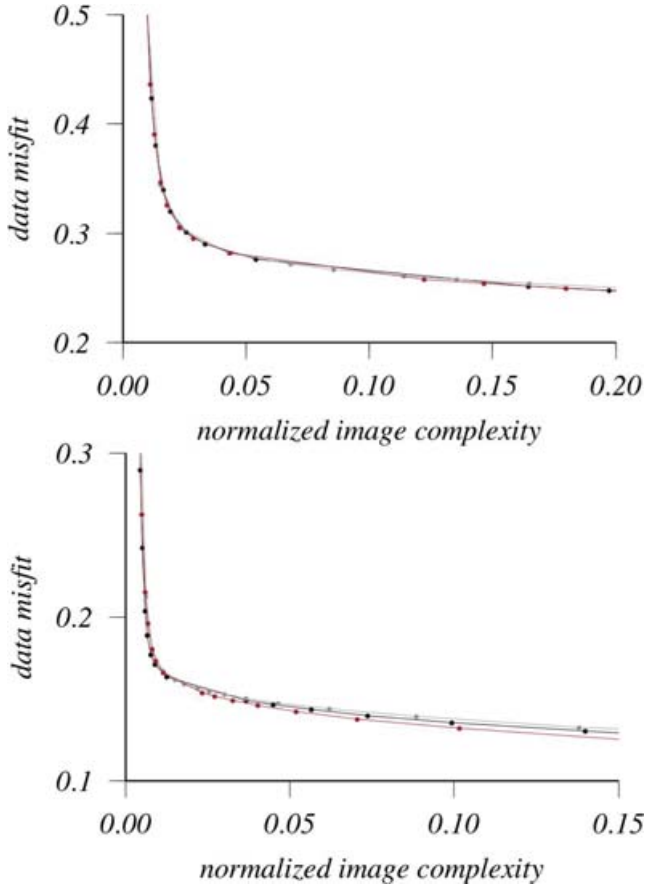

Figure 14. L-curve analysis results from analytical versus numerical kernels are compared: (top panel) same as Fig. 4, 150-s Love waves only, the red curve corresponds to numerical kernel inversions; (bottom panel) same as Fig. 8, that is, only data most sensitive to the high-resolution area are considered. The graph in the bottom panel is enlarged with respect to the top panel, to emphasize the slight downward shift of the numerical kernel L-curve (again in red).

is only partially reversed when the trade-off analysis is limited to the high-resolution region, and to phase-anomaly observations that are most sensitive to it (Section 4.3).

Perhaps the most likely explanation to this controversial result is proposed by Zhou *et al.* (2005), who noted that the phase-velocity inverse problem is complicated by the strongly oscillatory nature of the corresponding kernels (Figs 1 and 2): according to Zhou *et al.* (2005, fig. 16), oscillations in the sidebands of $K(\omega, \theta, \phi)$ result from the neglect of azimuthal dependence in surface wave scattering, required to describe the latter as a 2-D phenomenon (i.e. in terms of a set of phase-velocity maps) while it is, strictly speaking, 3-D. Zhou *et al.* (2005) inverted phase-anomaly data to derive 3-D maps of shear-velocity heterogeneity in the mantle and, accordingly, their trade-off analysis is in better agreement with theoretical expectations.

Another, potentially important limiting factor of scattering theory, as it is formulated here and in most of the literature (e.g. Spetzler *et al.* 2002; Montelli *et al.* 2004a,b; Zhou *et al.* 2004, 2005), is the far-field approximation on which it is grounded (Snieder 1987), resulting in $K(\omega, \theta, \phi)$ being singular at seismic source and receiver, and not strictly valid in their vicinity (e.g. Favier *et al.* 2004). At least one proof that $K(\omega, \theta, \phi)$, albeit singular, be integrable exists (Friederich 1999, appendix E), and is in agreement with the relative stability of the Born theoretical inverse problem. In Appendix B (Figs 14 and 15), I use numerical results from Peter *et al.* (2006) to drop the said far-field approximation, and repeat the L-curve and visual tomographic analyses. I conclude that the loss in model quality resulting from the far-field approximation is not relevant at the present level of resolution. It might become important, for example, for experiments aimed at resolving small-scale heterogeneities underlying a dense array of receivers.

A visual investigation of equivalently regularized phase-velocity maps (Section 6) shows (i) that both Born and ray theoretical results are in agreement with independent, regional-scale geophysical

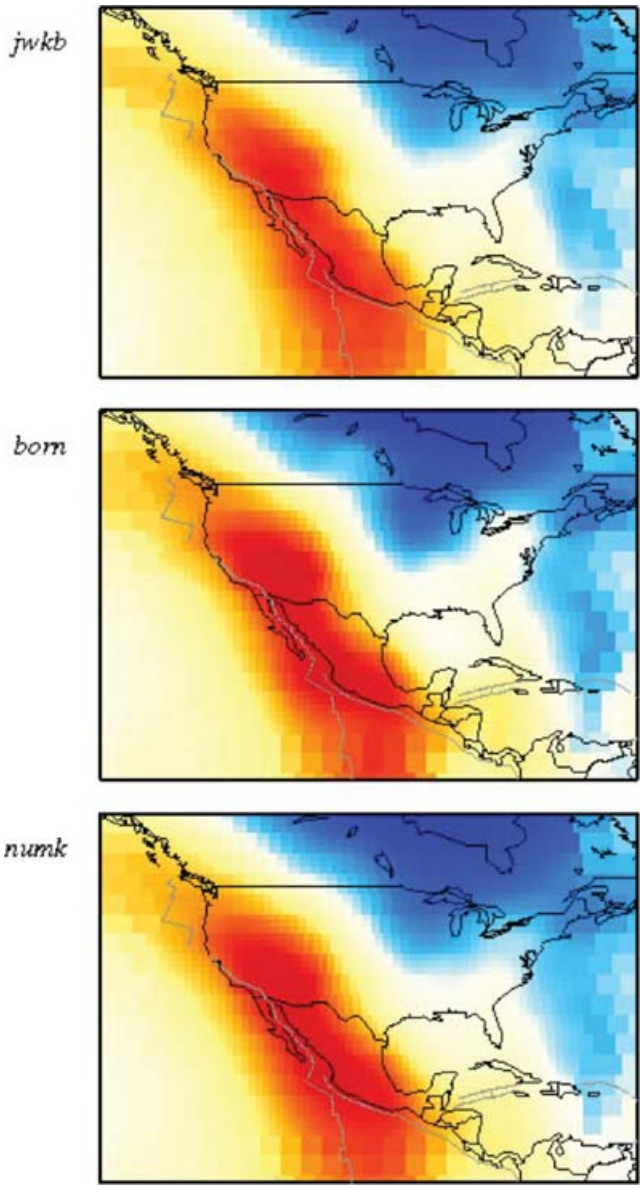

Figure 15. Phase velocity of 150-s Love waves, from ray-theory (top panel), analytical scattering-theory (middle panel) and numerical scattering-theory (bottom panel) formulations. Units and colour-scale are as in Figs 11 and 12.

observations; (ii) that, even at the regional scale, differences in the pattern of phase-velocity heterogeneity mapped via Born versus ray theory are negligible; (iii) that differences in amplitude are more visible, but still very small compared to perturbations in solution models resulting from small perturbations to the value of the damping parameter. Ray theory is still a valuable tool for surface wave phase-velocity tomography.

ACKNOWLEDGMENTS

This work would not have been possible without the kind help of Jesper Spetzler. Daniel Peter calculated numerical finite-frequency kernels (Appendix B). I am indebted to Göran Ekström for providing his database as well as many insightful comments. Thanks also to Domenico Giardini for his support and encouragement; Yann Capdeville, Jeannot Trampert and John Woodhouse for some clarifying discussions; Andrew Curtis, Barbara Romanowicz and one anonymous reviewer for their comments and suggestions; the SPICE (Seismic wave Propagation and Imaging in Complex media: a European network) project. All figures were done with GMT (Wessel & Smith 1991).

REFERENCES

Akaike, H., 1974. A new look at the statistical model identification, *IEEE Trans. Autom. Contr.,* **19,** 716–723.

Bender, C.M. & Orszag, S.A., 1978. *Advanced Mathematical Methods for Scientists and Engineers,* McGraw-Hill, New York.

Boschi, L., 2003. Measures of resolution in global body wave tomography, *Geophys. Res. Lett.,* **30,** 1978, doi:10.1029/2003GL018222.

Boschi, L. & Dziewonski, A.M., 1999. 'High' and 'low' resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *J. geophys. Res.,* **104,** 25 567–25 594.

Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface-wave phase velocity anomalies, *J. geophys. Res.,* **107,** doi:10.1029/2005GL025063.

Boschi, L., Becker, T.W., Soldati, G. & Dziewonski, A.M., 2006. On the relevance of Born theory in global seismic tomography, *Geophys. Res. Lett.,* **33,** L06302, doi:10.129/2000JB000059.

Capdeville, Y., 2005. An efficient Born normal mode method to compute sensitivity kernels and synthetic seismograms in the Earth, *Geophys. J. Int.,* **163,** 639–646.

Carannante, S. & Boschi, L., 2006. Databases of surface wave dispersion, *Ann. Geophys.,* **48,** 945–955.

Červeny, V., 1985. The application of ray-tracing to the propagation of shear waves in complex media, in *Seismic Shear Waves. Part A: Theory,* ed. Dohr, G., Geophysical Press, Amsterdam.

Clévédé, E. & Lognonné, P., 1996. Fréchet derivatives of coupled seismograms with respect to an anelastic rotating Earth, *Geophys. J. Int.,* **124,** 456–482.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press, Princeton, New Jersey.

Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes–I. Theory, *Geophys. J. Int.,* **141,** 151–174.

Ekström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.,* **102,** 8137–8157.

Favier, N., Chevrot, S. & Komatitsch, D., 2004. Near-field influence on shear wave splitting and traveltime sensitivity kernels, *Geophys. J. Int.,* **156,** 467–482.

Friederich, W., 1999. Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering, *Geophys. J. Int.,* **136,** 180–204.

Godey, S., 2002. Structure of the Uppermost Mantle Beneath North America: Regional Surface Wave Tomography and Thermo-chemical Interpretation, *PhD thesis*, Utrecht University.

Godey, S., Snieder, R., Villase nor, A. & Benz, H.M., 2003. Surface wave tomography of North America and the Caribbean using global and regional broad-band networks: phase velocity maps and limitations of ray theory, *Geophys. J. Int.,* **152,** 620–632.

Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the L-curve, *Siam Rev.,* **34,** 561–580.

Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes–II. Examples, *Geophys. J. Int.,* **141,** 175–203.

Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y., 1990. Whole mantle *P*-wave travel time tomography, *Phys. Earth planet. Inter.,* **59,** 294–398.

Komatitsch, D., Ritsema, J. & Tromp, J., 2002. The spectral-element method, Beowulf computing and global seismology, *Science,* **298,** 1737–1742.

Li, X.-D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121,** 695–709.

Li, X.-D. & Romanowicz, B., 1996. Global mantle shear-velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.,* **101,** 22 245–22 272.

Li, X.-D. & Tanimoto, T., 1993. Waveforms of long-period body waves in a slightly aspherical Earth model, *Geophys. J. Int.,* **112,** 92–102.

Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R. & S.-H. 2004a. Hung, Finite-frequency tomography reveals a variety of plumes in the mantle, *Science,* **303,** 338–343.

Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.-H., 2004b. Global *P* and *PP* traveltime tomography: rays versus waves, *Geophys. J. Int.,* **158,** 637–654, doi: 10.1111/j.1365-246X.2004.02346.x.

Nataf, H.-C. & Ricard, Y., 1996. 3SMAC: an *a priori* tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth planet. Inter.,* **95,** 101–122.

Nerem, R.S. et al., 1994. Gravity model development for TOPEX/Poseidon: joint gravity models 1 and 2, *J. geophys. Res.,* **99,** 24 421–24 447.

Nerem, R.S., Jekeli, C. & Kaula, W.M., 1995. Gravity field determination and characteristics: retrospective and prospective, *J. geophys. Res.,* **100,** 15 053–15 074.

Perry, H.K.C., Forte, A.M. & Eaton, D.W.S., 2003. Upper-mantle thermochemical structure below North America from seismic-geodynamic flow models, *Geophys. J. Int.,* **154,** 279–299.

Peter, D. & Boschi, L., 2005. Membrane waves and finite-frequency effects in surface wave tomography, *Geophys. Res. Abstr.,* **7,** 931, SRef-ID:1607-7962/gra/EGU05-A-00931.

Peter, D., Tape, C.H., Boschi, L. & Woodhouse, J.H., 2006. Surface wave tomography: membrane waves and adjoint methods, *Geophys. J. Int.,* submitted.

Ritzwoller, M.H., Shapiro, N.M., Barmin, M.P. & Levshin, A.L., 2002. Global surface wave diffraction tomography, *J. geophys. Res.,* **107,** 2335, doi:10.1029/2002JB001777.

Romanowicz, B., 1987. Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the Earth's normal modes, *Geophys. J. R. astr. Soc.,* **90,** 75–100.

Sieminski, A., Lévêque, J.-J. & Debayle, E., 2004. Can finite-frequency effects be accounted for in ray theory surface wave tomography? *Geophys. Res. Lett.,* **31,** doi:10.1029/2004GL021402.

Simons, M. & Hager, B.H., 1997. Localization of the gravity field and the signature of glacial rebound, *Nature,* **390,** 500–504.

Snieder, R., 1986a. 3-D linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. R. astr. Soc.,* **84,** 581–605.

Snieder, R., 1986b. The influence of topography on the propagation and scattering of surface waves, *Phys. Earth planet. Inter.,* **44,** 226–241.

Snieder, R., 1987. Surface Wave Scattering Theory, *PhD thesis*, Utrecht University.

Snieder, R., 1988. Large-scale waveform inversions of surface waves for lateral heterogeneity 1. Theory and numerical examples, *J. geophys. Res.,* **93,** 12 055–12 065.

Snieder, R. & Nolet, G., 1987. Linearized scattering of surface waves on a spherical Earth, *J. Geophys.,* **61,** 55–63.

Soldati, G. & Boschi, L., 2005. The resolution of whole Earth seismic tomographic models, *Geophys. J. Int.,* **161,** 143–153.

Spetzler, J., Trampert, J. & Snieder, R., 2001. Are we exceeding the limits of the great circle approximation in surface wave tomography?, *Geophys. Res. Lett.,* **28,** 2341–2344.

Spetzler, J., Trampert, J. & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.,* **149,** 755–767.

Tanimoto, T., 1990. Modelling curved surface wave paths: membrane surface wave synthetics. *Geophys. J. Int.,* **102,** 89–100.

Tanimoto, T., 2003. Geometrical approach to surface wave finite frequency effects, *Geophys. Res. Lett.,* **30,** 1993, doi:10.1029/2003GL017475.

Tape, C.H., 2003. Waves on a Spherical Membrane, *MSc thesis*, Oxford University.

Tromp, J., Tape, C. & Liu, Q.Y., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Weisstein, E.W., 2003. *CRC Concise Encyclopedia of Mathematics,* Chapman & Hall/CRC, Boca Raton, Florida.

Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72,** 445–446.

Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized earth model, *Geophys. J. R. astr. Soc.,* **68,** 653–673.

Yoshizawa, K. & Kennett, B.L.N., 2005. Sensitivity kernels for finite-frequency surface waves, *Geophys. J. Int.,* **162,** 910–926, doi: 10.1111/j.1365-246X.2005.02707.x.

Zhou, Y., Dahlen, F.A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.,* **158,** 142–168, doi: 10.1111/j.1365-246X.2004.02324.x.

Zhou, Y., Dahlen, F.A., Nolet, G. & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.,* **163,** 1087–1111, doi: 10.1111/j.1365-246X.2005.02780.x.

APPENDIX A: SPHERICAL HARMONIC PARAMETRIZATION

When phase velocity is described as a linear combination of spherical harmonics (Spetzler *et al.* 2002), eq. (8) is replaced by

$$A_{ij} = \int_\Omega K_{i\text{th datum}}(\omega, \theta, \phi) Y_{lm}(\theta, \phi)\, d\Omega, \tag{A1}$$

where $Y_{lm}(\theta, \phi)$ is the degree l, order m, real spherical harmonic (e.g. Dahlen & Tromp 1998, appendix B), and a one-to-one correspondence is defined between values of j (column-index of $\mathbf{A}$) and l, m couples.

As long as spherical harmonics are normalized, this is equivalent to stating that A_{ij} equals the l, m coefficient of a spherical harmonic expansion of $K_{i\text{th datum}}(\omega, \theta, \phi)$.

It is then convenient to first find (by numerical integration) the harmonic coefficients of $K(\omega, \theta, \phi)$ for a source at (0°, 0°) and stations located along the equator, and later, for each datum, rotate to the proper source–station geometry the kernel associated with the same epicentral distance. This procedure is analogous to the one I have followed with my pixel parametrization, but in the case of spherical harmonics it is particularly simple, as $K(\omega, \theta, \phi)$ can be rotated by a simple dot-multiplication between a quickly determined rotation matrix, and the vector of $K(\omega, \theta, \phi)$'s harmonic coefficients (e.g. Dahlen & Tromp 1998, section C.8.6).

Despite its elegance, the spherical harmonic parametrization might not be adequate to represent features of high spatial frequency, unless very high values of l are considered. Spetzler *et al.* (2002) employed a spherical harmonic expansion up to degree 40; I follow their procedure to find the corresponding harmonic coefficients of $K(\omega, \theta, \phi)$, and I show in Fig. 13 the spherical-harmonic version of some of the Fréchet kernels illustrated in Fig. 1; it is clear that a degree-40 harmonic expansion is not sufficient to represent these very heterogeneous functions; unless higher-degree coefficients of $K(\omega, \theta, \phi)$ are accounted for, a certain loss of resolution is unavoidable.

This issue is particularly important in the context of Born theoretical tomography, where, assuming that the data coverage be adequate, resolution should not be limited by the wavelength of inverted observations (Spetzler *et al.* 2001, 2002; Sieminski *et al.* 2004).

My pixelized parametrization (see Fig. 2) is a way to avoid the problem altogether.

APPENDIX B: NUMERICAL SENSITIVITY KERNELS

Peter *et al.* (2006) found phase-velocity sensitivity kernels for isolated surface wave modes, applying an adjoint approach similar to that of Tromp *et al.* (2005) to a finite-difference membrane-wave algorithm (Tanimoto 1990; Tape 2003). Like Spetzler *et al.*'s (2002) approach, Peter *et al.*'s (2006) approach is only strictly valid in the neglect of mode coupling and in the assumption that lateral heterogeneity in upper mantle structure be smooth. Peter *et al.*'s (2006) numerical kernels differ, however, from Spetzler *et al.*'s (2002) ones, in that they are non-zero also for perturbations in phase velocity at azimuths $>\pi/2$ with respect to the source–station great circle, which Spetzler *et al.* (2002) neglected; more importantly, they differed from most analytically calculated sensitivity kernels (Li & Tanimoto 1993; Li & Romanowicz 1995 1996; Dahlen *et al.* 2000, and following publications by the Princeton group; Yoshizawa & Kennett 2005) in that they are not grounded upon a far-field approximation, and should be as valid in the vicinity of source and receiver as they are away from them.

Limiting myself to Love waves at 150 s, I conduct an L-curve analysis (as in Section 4) of Born theoretical inversions based on Peter *et al.*'s (2006) numerical sensitivity kernels, and show the results in Fig. 14. At the global scale (top panel), the new L-curve does not visibly differ from those at the bottom left-hand panel of Fig. 4 (reproduced in Fig. 14 for convenience). If, as in Section 4.3, only data most sensitive to structure underlying North America are considered (bottom panel), the L-curve resulting from numerical sensitivity kernels lies everywhere below the others; the difference is, however, barely detectable, and still very small compared to the theoretical prediction of Fig. 6.

After selecting a preferred numerical kernel solution model with the criterion described in Section 4.1, I plot it in Fig. 15, accompanied by the corresponding Born and ray theoretical tomographic results from Fig. 11. Differences in pattern and amplitude of heterogeneities found from numerical versus analytical sensitivity kernels are certainly not larger than those found between Born and ray theoretical results, that is, they are negligible. This is in agreement with Hung *et al.*'s (2000) comparison of analytical (far-field) and numerical kernels, and with Friederich's (1999, appendix E) proof that the far-field kernels' singularity at source and receiver is integrable.